

Universidade de Brasília
Instituto de Geociências

**IV CURSO DE ESPECIALIZAÇÃO EM GEOPROCESSAMENTO
AMBIENTAL**

**Reflectância acumulada e mineração de dados:
ferramentas conjuntas para a classificação orientada a
objeto de área de cerrado antropizado**

THALLITA OLIVEIRA DE GRANDE

Monografia de conclusão de
curso de especialização

ORIENTADORA: Profa. Dra. Tati Almeida (IG / UnB)

BRASÍLIA- DF

2016



Universidade de Brasília
Instituto de Geociências

**IV CURSO DE ESPECIALIZAÇÃO EM GEOPROCESSAMENTO
AMBIENTAL**

**Reflectância acumulada e mineração de dados:
ferramentas conjuntas para a classificação orientada a
objeto de área de cerrado antropizado**

Thallita Oliveira de Grande

Monografia de conclusão de
curso de especialização

ORIENTADORA: Profa. Dra. Tati Almeida (IG / UnB)

BRASÍLIA- DF

2016



Universidade de Brasília
Instituto de Geociências

Reflectância acumulada e mineração de dados: ferramentas conjuntas para a classificação orientada a objeto de área de cerrado antropizado

Thallita Oliveira de Grande

Monografia de conclusão de
curso de especialização

ORIENTADORA: Profa. Dra. Tati Almeida (IG / UnB)

COMISSÃO JULGADORA:

Prof. Dr. Gustavo Macedo de Mello Baptista

Prof. Dr. Edson Eyji Sano

BRASÍLIA- DF

2016

AGRADECIMENTOS

Agradeço ao Instituto de Geociências (IG), pela oportunidade de crescimento e ganho de aprendizado por meio deste enriquecedor e excelente curso de geoprocessamento ambiental. Em especial, sou grata à orientação generosa e paciente da professora Dra. Tati Almeida, uma professora que é grande inspiração para todos. Agradeço também aos professores Dr. Edson Sano e Dra. Rejane Cicerelli, e aos colegas Rodrigo Antunes e Denilson Passos, pelas dicas e discussões, que muito contribuíram para a concretização do trabalho. Ao Sérgio Paulo Brito (Serjão), por organizar e facilitar as questões administrativas do curso. A ESRI, pela disponibilização do pacote de ferramentas que compõem a família ArcGis 10.x por intermédio do contrato nº 2011 MLK 8733, e a IMAGEM pelo apoio e viabilidade da concretização do termo de uso entre o IG e a ESRI, e pelo suporte aos softwares.

SUMÁRIO

ABSTRACT / RESUMO.....	06
INTRODUÇÃO.....	07
MÉTODOS.....	09
RESULTADOS E DISCUSSÃO.....	13
CONCLUSÃO.....	28
REFERÊNCIAS BIBLIOGRÁFICAS.....	28

REFLECTÂNCIA ACUMULADA E MINERAÇÃO DE DADOS: FERRAMENTAS CONJUNTAS PARA A CLASSIFICAÇÃO ORIENTADA A OBJETO DE ÁREA DE CERRADO ANTROPIZADO

Grande, T. O.^{1*} & Almeida, T.²

^{1*} Programa de Pós Graduação em Ecologia, Universidade de Brasília, Brasil. tata9go@hotmail.com

² Instituto de Geociências, Universidade de Brasília, Brasil.

ABSTRACT *The Brazilian Savannah (Cerrado) has been intensely replaced by agriculture in recent decades, requiring a constant effective and automated monitoring of the biome. Associate human knowledge with automatic interpretation of images such as provided by the object-oriented classification and data mining overcomes the traditional classifications of land use and land cover. Furthermore, the use of accumulated reflectance of multi-temporal images increases the Cerrado class discrimination. We tested the association of cumulative reflectance technique as a basis for data mining and object-oriented classification to classify the Parque Nacional das Emas - GO and surroundings. We use four Landsat 8 scenes in format of reflectance of the Earth's surface. We define six classes: agriculture, livestock, wetland, savannah, forest and grassland. We gradually accumulate the reflectance images and use the image with the four cumulative dates. We use the endmembers methodology to eliminate the noises of the image. Using the InterIMAGE software, we segmented the image, extracting spectral attributes of some samples and applying two combinations of attributes ("mean + mode", "all attributes") for data mining in the Weka software. The decision tree generated in Weka based the decision rules for classification in InterIMAGE. The distinction between classes increased with the temporal accumulation of reflectance for all bands, among all classes and dates accumulation levels, and the best image was the one referring to the sum of the four dates. The classification based on the attributes association "mean + mode" had no problems in relation to the decision rules processing, unlike the association "all attributes", which the livestock class had not run. A posteriori tests with more samples and attributes also failed in running. The "mean + mode" classification showed good accuracy (Overall Accuracy: 69%; Kappa: 58%; TAU: 63%). The association of these three techniques have a good potential to differentiate native from anthropic Cerrado vegetation. We however recommend closer examination of the classification object-oriented in InterIMAGE software, its computing capacity and the thresholds of the number of attributes, the complexity of trees and decision rules established for the classification.*

Keywords: *multitemporal analysis, remote sensing, human disturbance, data mining, OBIA, segmentation*

RESUMO O Cerrado vem sendo intensamente substituído nas últimas décadas pela agropecuária, tornando constante a necessidade de monitorar o bioma de forma eficaz e automatizada. Associar o conhecimento humano com a interpretação automática de imagens, como por meio da classificação orientada a objeto e mineração de dados, supera as classificações tradicionais do uso e cobertura da terra. Ademais, o uso da reflectância acumulada de imagens multitemporais aumenta a discriminação dos alvos do Cerrado. Testamos a associação da técnica de reflectância acumulada como base para a mineração de dados e classificação orientada a objeto, para classificar a região do Parque Nacional das Emas - GO, e adjacências. Usamos quatro cenas Landsat 8 de reflectância na superfície terrestre. Definimos seis classes: agricultura, pecuária, várzea, savana, floresta e campo. Acumulamos progressivamente as imagens de reflectância e usamos a imagem com as quatro datas acumuladas. Usamos a metodologia de *endmembers* para eliminar os ruídos da imagem. No InterIMAGE,

segmentamos a imagem, extraímos atributos espectrais de algumas amostras, e aplicamos duas combinações de atributos (“média + moda”, “todos atributos”) na mineração de dados, no programa Weka. A árvore de decisão gerada no Weka subsidiou as regras de decisão para a classificação no InterIMAGE. A diferenciação entre alvos aumentou com a acumulação temporal da reflectância, em todas as bandas, entre todas as classes e níveis de acumulação das datas, sendo a melhor imagem aquela do somatório das quatro datas. A classificação baseada na associação de atributos “média + moda” não teve impedimentos no processamento das regras de decisão, diferentemente da associação de “todos os atributos”, cuja classe “pecuária” não foi executada. Testes a posteriori com mais amostras e atributos também não conseguiram ser executados pelo programa. A classificação “média + moda” apresentou boa acurácia (Exatidão Global: 69%; *Kappa*: 58%; TAU: 63%). As associações destas três técnicas têm bom potencial na diferenciação de vegetação nativas e antrópicos do Cerrado. Recomendamos, contudo, investigações mais detalhadas sobre a classificação orientada a objeto no programa InterIMAGE, sua capacidade computacional e os limites da quantidade de atributos, da complexidade das árvores e regras de decisão implantadas para a classificação.

Palavras-chave: análise multitemporal, sensoriamento remoto, antropização, data mining, OBIA, segmentação

INTRODUÇÃO

O Cerrado, segundo maior bioma brasileiro e um dos *hotspot* de biodiversidade mundial (Myers et al. 2000), é um bioma constituído, originalmente, por um mosaico natural de ambientes, como os savânicos e campestres (em maior proporção) e os florestais (em menor proporção) (Sano et al. 2010). Apesar da grande diversidade de ambientes, o Cerrado vem sofrendo forte alteração do seu mosaico, uma vez que é, atualmente, o bioma brasileiro submetido a maior pressão do desmatamento, fundamentalmente pela expansão da fronteira do agronegócio (Klink e Machado 2005; Faleiro et al. 2013). Em toda sua extensão espacial, suas fitofisnomias vêm sendo substituídas intensamente pela agricultura e pecuária, os dois tipos dominantes do uso da terra (Sano et al. 2010).

Diante desse cenário de rápida e intensa mudança espacial no Cerrado, conhecer a dinâmica desta transformação e, sobretudo, ter acesso rápido e preciso a mesma por meio de levantamentos eficientes do uso e cobertura da terra, é fundamental (Sano et al. 2007). A ferramenta do sensoriamento remoto possibilita que isto seja feito, e os padrões de organização do espaço sejam descritos e compreendidos (Brannstrom et al. 2008). Sob a ainda atual limitação de acesso às imagens de alta resolução espectral e espacial, o uso de imagens Landsat continua sendo um canal de informações remotas acessíveis e factível. Por isso, aprimorar técnicas e protocolos aplicáveis de monitoramento da cobertura da terra utilizando imagens Landsat é algo relevante em termos científicos e, sobretudo, de política de proteção e monitoramento (Nascimento e Sano 2010). No contexto do Cerrado, entretanto, existem particularidades que se configuram como barreiras aos

mapeamentos de uso e cobertura da terra e, conseqüentemente, à transformação dessas informações em ações políticas práticas (Brannstrom et al. 2008).

Por exemplo, no Cerrado há uma natural heterogeneidade, que forma um gradiente espacial entre fitofisionomias campestres e florestais (Pinheiro e Durigan 2009), cuja separabilidade não é algo trivial de ser alcançado, uma vez que as diferentes formações não se distinguem espacialmente de forma categórica e possuem, além disso, respostas espectrais semelhantes (Sano et al. 2009; Costa 2014). O avanço do agronegócio sobre o bioma imprime na paisagem a quebra desta heterogeneidade em substituição à uma homogeneidade espacial, onde classes como pastagens cultivadas e plantações também apresentam respostas espectrais confundidas com as das formações nativas (Sano et al. 2009). Estas características impõem ao contexto do Cerrado o desafio permanente da automação do processo de classificação, de modo que o mapeamento da paisagem alcance objetividade e abarque grandes volumes de dados (DeFries e Chan 2000).

Diante destes desafios, a abordagem da classificação orientada a objeto associa o conhecimento trazido pelo operador especialista com a interpretação automática de imagens de sensoriamento remoto (Novak 2009). Nela, a distinção entre alvos não se resume às características espectrais dos pixels, como usado pelos classificadores tradicionais, algo pouco eficiente no reconhecimento automático do padrão de alvos. Ao invés disso, esta usa como unidade de processamento objetos segmentados com base em características comuns (forma, textura, espectro, vizinhança, geometria, etc.), e incorpora estes atributos na classificação, agrupando os objetos que possuem uma gama de atributos semelhantes (Liu et al. 2006; Blaschke 2010). Logo, o uso da abordagem orientada a objeto na análise de paisagens preenchidas de alvos heterogêneos e complexos, como é o caso do Cerrado antropizado, parece mais adequado do que o uso de classificadores automáticos convencionais (Novak 2009).

Dada a grande quantidade de dados que podem ser gerados na definição dos atributos caracterizadores dos segmentos na classificação orientada a objeto, é útil o uso de técnicas de Mineração de Dados, definidas como algoritmos computacionais que extraem novas informações de grandes bases de dados. Estes algoritmos são capazes de expressar padrões embutidos nos dados, os quais podem ser utilizados na classificação (Whitten e Frank 2005; Hall et al. 2009). Por exemplo, as Árvores de Decisão são algoritmos comumente usados, que geram definições em formato de árvore (estrutura hierárquica de regras com nós e ramificações), onde cada nó representa uma decisão sobre um atributo. Esta estrutura pode, então, ser aplicada na definição de regras de decisão das classes de alvos, sobre toda a base de dados, no processo de classificação (Quinlan 1993).

Por fim, considerando a aplicabilidade do uso de imagens Landsat no mapeamento do uso e cobertura da terra no Cerrado, estudos mostram que o uso de imagens Landsat multitemporais, em detrimento das monotemporais, é algo vantajoso na classificação do bioma, uma vez que o acúmulo

de aspectos espectrais das imagens, como sua refletância, ao longo de um intervalo temporal determinado, aumenta a capacidade de discriminação entre os alvos (Nascimento e Sano 2010; Hermuche e Sano 2011).

Nosso objetivo foi, então, utilizar a técnica de refletância acumulada como base para a aplicação das técnicas de mineração de dados (*data mining*), seguida da classificação orientada a objeto, para fins da diferenciação e classificação de fitofisionomias nativas e formações agropecuárias do Cerrado, utilizando imagens Landsat. Partimos da premissa de que em imagens com refletância acumulada a diferença entre alvos se torna mais acentuada. Assim, nossa hipótese era a de que as técnicas de mineração de dados, capaz de captar os limiares sutis de diferenciação de alvos, juntamente com a segmentação orientada a objeto, que associa diversos atributos dos alvos na sua diferenciação, aplicadas sobre imagens com refletância acumulada, são uma associação eficaz na distinção das fitofisionomias nativas e formações agropecuárias do Cerrado, cujas respostas espectrais são semelhantes. Especificamente, testamos se a refletância acumulada difere significativamente entre os níveis de acumulação das datas (e quais) e entre as classes de uso e cobertura da terra.

MÉTODOS

Área de estudo

A área de estudo é um recorte que compreende a quase totalidade da área do Parque Nacional das Emas (18°07'17" S, 52°54'54" W), no sudoeste goiano, e suas adjacências, dominadas por áreas agropecuárias, totalizando 176.781,42 hectares. Escolhemos esta região em função do prévio conhecimento das classes de uso e cobertura da terra presentes na localidade, dado tanto pela experiência dos autores como pela caracterização da região disponível na literatura científica (Ibama 2004).

Imagens de Sensoriamento Remoto

Utilizamos quatro cenas de imagens Landsat 8 OLI/TIRS de refletância na superfície terrestre ortorretificadas, disponíveis no site do *U.S. Geological Service* (<http://earthexplorer.usgs.gov/>). Dentro do intervalo temporal entre os anos de 2013 e 2014, escolhemos cenas com menos de 10% de cobertura de nuvens (Órbita: 224; Ponto: 73; datas: 18/07/13 – D1, 03/08/13 – D2, 19/08/13 – D3, 11/02/14 – D4), o que limitou a escolha às quatro cenas supracitadas. Determinamos esse intervalo em função da oferta das imagens Landsat 8 de refletância na superfície terrestre, disponíveis apenas a partir de 2013, e da necessidade de adotar um intervalo igual ou superior às imagens que poderiam ser usadas para a verificação da qualidade da classificação (vide seção).

Reflectância acumulada

No programa ENVI 5.1, para cada cena Landsat, unimos as sete bandas em um único arquivo e recortamos a área de interesse. Somamos, por banda, as cenas das quatro datas (D1, D2, D3, D4) de maneira acumulada, de modo que $A1 = D1$, $A2 = D1+D2$, $A3 = D1+D2+D3$, $A4 = D1+D2+D3+D4$. Com base em classificações prévias da área (Ibama 2004) e na experiência pessoal dos autores, escolhemos visualmente na paisagem uma área de interesse (ROI) de tamanho fixo ($5 \times 5 = 25$ pixels) por classe, cujo tamanho julgamos ser suficiente para amostrar as características da reflectância das mesmas. As classes eram: agricultura, pecuária, várzea, savana, floresta e campo. Obtivemos a média da reflectância dos ROIs para cada classe, tanto por banda e quanto pelas datas individuais e acumuladas. Aplicamos, por bandas, ANOVAs *two-way* sem repetição para testar se a reflectância diferia significativamente entre os níveis de acumulação das datas e entre as classes, e testes t de *Student* uni-caudais pareados para testar entre quais níveis de acumulação das datas as reflectâncias acumuladas diferiam mais significativamente entre si.

Pré-processamento de imagem

Aplicamos sobre a imagem de reflectância acumulada das quatro datas a metodologia de *endmembers*, tal qual utilizado por Hermuche e Sano (2010), para eliminar os ruídos intrínsecos às imagens, para fins de posterior classificação. Como explica Carvalho Júnior et al. (2003 e 2007), esta metodologia consiste nas seguintes etapas: 1) redução da dimensão espectral da imagem (*Minimization Noise Fraction* – MNF) (Green et al. 1988; Boardman e Kruse 1994), técnica derivada da Análise de Componentes Principais (PCA), em que criam-se eixos (componentes principais) ordenados conforme a intensidade do ruído da cena; 2) redução da dimensão espacial da imagem (*Pixel Purity Index* - PPI), técnica que identifica pixels espectralmente puros na imagem, ou seja, aqueles que se localizam nos vértices de um diagrama de dispersão (escaterograma n-dimensional) dos componentes principais (MNFs); 3) projeção espacial dos dados em um escaterograma n-dimensional, que possibilita a visualização de nuvens de pontos dos componentes principais. Com o programa ENVI 5.1, utilizamos a ferramenta *Spectral Hourglass Wizard* para a aplicação das técnicas supracitadas. Aplicamos 5.000 iterações (quantidade suficiente para a estabilização do processamento) e reduzimos para três a dimensionalidade dos dados, uma vez que os três primeiros MNFs (componentes) foram responsáveis pela maior variação dos dados (Figura 1); usamos para os demais parâmetros o padrão (*default*) do programa. Finalmente, usando a ferramenta *Inverse MNF Rotation*, invertemos a imagem do MNF de três dimensões, redistribuindo os espectros sem ruídos para sua dimensão espectral original (7 bandas) (Carvalho et al. 2002; Hermuche e Sano 2010).

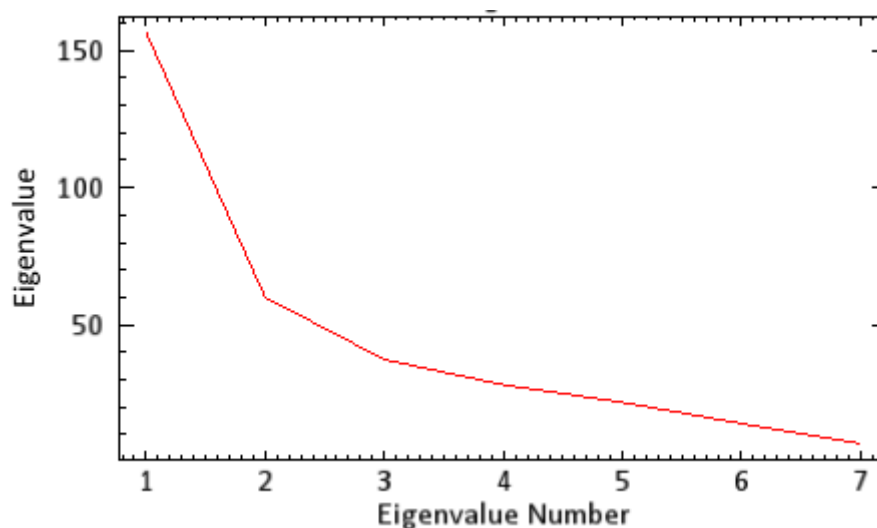


Figura 1. Gráfico dos autovalores resultantes da etapa de redução da fração ruidosa (*Minimization Noise Fraction – MNF*) da técnica *endmembers*, indicando o ponto de inflexão na terceira MNF (*eigenvalue number*), antes do qual concentra-se a maior parte do sinal da imagem e a partir do qual concentra-se o ruído.

Segmentação, mineração de dados e classificação orientada a objeto

Utilizamos a imagem do somatório das quatro datas processada em três dimensões e redistribuída em sete bandas como o dado de entrada no programa InterIMAGE 4.2. Segmentamos a imagem de acordo com os parâmetros descritos na tabela 1 (Antunes et al. 2015), e utilizando o segmentador proposto por Baatz e Schäpe (2000) (Passo et al. 2013) escolhemos um total de 82 amostras de treinamento distribuídas entre as seis classes. Buscamos escolher amostras em locais que pudessem representar as classes com a maior fidelidade possível.

Como nossa hipótese previa a utilização de uma imagem com reflectância acumulada como base para as demais análises, decidimos extrair da mesma alguns atributos espectrais que julgamos ser mais relevantes para refletir o possível efeito da acumulação da reflectância nas análises, dentro de uma gama de atributos de variadas naturezas disponíveis pelo programa InterIMAGE. Para tal, extraímos das amostras, por banda e por classe, os seguintes atributos espectrais: média (valor médio de pixels), pixmax (valor de pixel máximo por banda), pixmin (valor de pixel mínimo por banda), amplitude (diferença entre os valores máximos e mínimos de pixel por banda), moda (valor mais frequente de pixels por banda), soma (somatório dos valores de pixel por banda). No programa Weka 3.6, realizamos a mineração dos dados supracitados, considerando, para cada atributo separadamente, as amostras de todas as classes juntas. De acordo com Antunes et al. (2015), utilizamos o algoritmo J48 e obtivemos, por atributo, uma árvore de decisão e suas respectivas estatísticas. Consideramos os melhores atributos aqueles que ofereciam o menor número de caminhos para a identificação das classes, a árvore com o menor tamanho e número de ramos, mais amostras classificadas corretamente

e menos classificadas incorretamente e um maior índice *Kappa*. Repetimos, então, a mineração de dados para as associações dos atributos “média + moda” e de “todos os atributos juntos”.

Realizamos, por fim, a classificação *TopDown* orientada a objeto da imagem de reflectância acumulada no programa InterIMAGE 4.2 (Baatz e Schäpe 2000; Passo et al. 2013), com base nos parâmetros descritos na tabela 1. Construímos as regras de decisão de cada nó (classes) da rede semântica com base na árvore de decisão gerada pela mineração de dados da associação de atributos “média + moda”, uma vez que o programa não conseguiu executar as regras de decisão da classe “pecuária” gerada pela associação de “todos os atributos juntos”, o que será discutido posteriormente.

Tabela 1. Parâmetros utilizados na segmentação para a coleta de amostras e na classificação orientada a objeto da imagem Landsat 8 de reflectância acumulada, do Parque Nacional das Emas e adjacências - Goiás.

Segmentação para coleta de amostras										
Operador TopDown	Bandas entrada	Peso bandas	Compacidade	Cor	Escala	Otimização	Confiabilidade	Distância Euclidiana	Operador BottomUp*	Ponto de parada
TA Baatz Segmenter	0,1,2,3, 4,5,6	1,1,1, 1,1,1,1	0.5	0.5	1000	No	0.2	20	Dummy BottomUp	None
Classificação										
TA Baatz Segmenter	0,1,2,3, 4,5,6	1,1,1, 1,1,1,1	0.5	0.5	100	No	0.2	20	Dummy BottomUp	None

* Demais parâmetros foram utilizados de acordo com o padrão (*default*) do programa: Regra de decisão *BottomUp*: All - *Spatial Resolve*; *Multi classes TopDown*: desmarcado.

Verificação da qualidade da classificação

Para fins de verificação da precisão da classificação orientada a objeto, seguimos a metodologia proposta por Congalton e Green (1999), em que uma matriz de confusão (quadrada) é construída com as amostras que foram atribuídas às classes do mapa a partir de duas classificações (a real e a de referência/verdade de campo). De acordo com o método, um número de pontos amostrais de verificação da exatidão da classificação é determinado pelo valor de uma distribuição qui-quadrado, seu grau de liberdade, o número de classes, um nível de confiança, a proporção da classe com maior área no mapa e um erro máximo. Consideramos o erro admissível de 0,1 e obtivemos um $N = 126$ de pontos amostrais. Geramos estes pontos amostrais aleatoriamente sobre a imagem utilizando a ferramenta *Random Selection (Hawths Analysis Tools)* do programa ArcGIS 10.2 e com uma verificação cuidadosa atribuímos a cada ponto sua respectiva classe. Em casos de dúvidas, conferimos tal verificação utilizando cenas de imagens *RapidEye* do ano de 2013 e do *Google Earth* do ano de 2014. Utilizamos estes dados como dados de referência (verdade de campo) para a construção da matriz de confusão. Finalmente, aplicamos três dos principais coeficientes de concordância para avaliação da acurácia da classificação: índice de Exatidão Global (Po) (Hellden e Stern 1980), *Kappa* (Cohen 1960) e TAU (Klecka 1980).

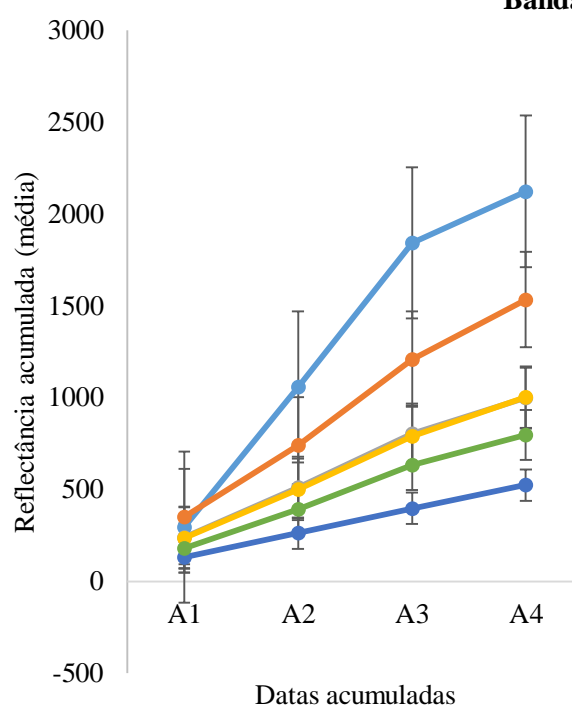
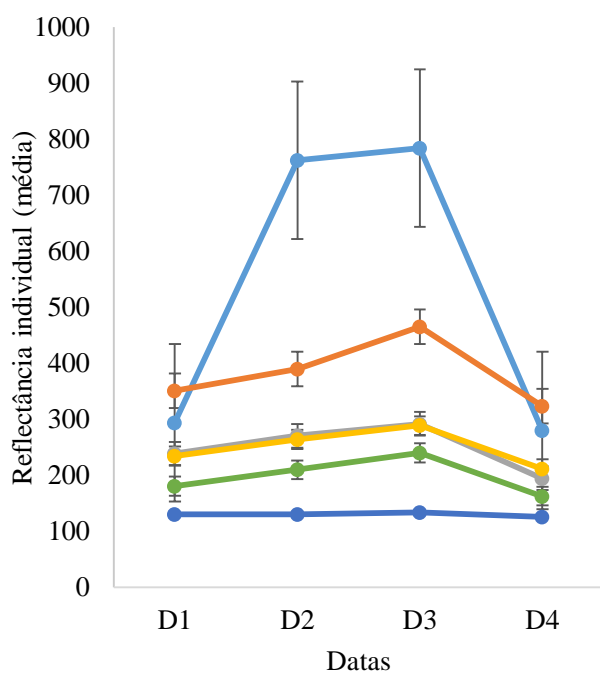
RESULTADOS E DISCUSSÃO

Reflectância acumulada

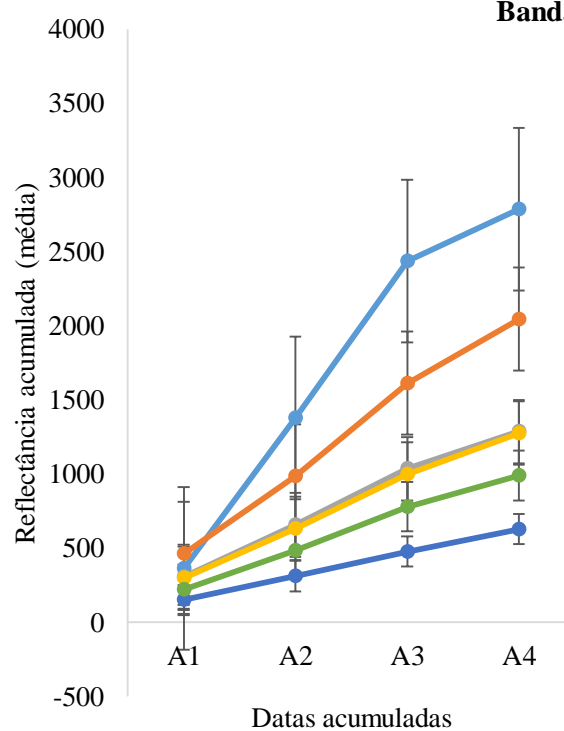
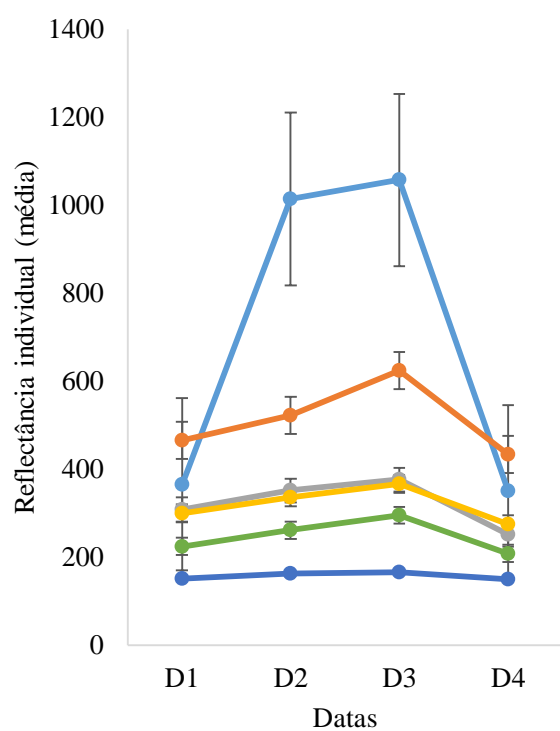
Por meio da figura 2 e das significâncias estatísticas das ANOVAs (Tabela 2), podemos observar que a técnica de acumulação da reflectância demonstrou, de fato, ser capaz de aumentar a diferenciação entre alvos, em todas as bandas, tanto entre as classes quanto entre os níveis de acumulação das datas, o que corroborou a premissa que assumimos. Para todas as bandas, o melhor nível de acumulação foi o que resultou do somatório das quatro datas. Como pontuado por Nascimento e Sano (2010), quanto mais cenas consideradas, maior é a capacidade de discriminação entre alvos. Pudemos confirmar isto observando as comparações entre os níveis de acumulação. Entre os níveis mais extremos de acumulação de datas, como o par A4-A1, houve maior valor do teste estatístico e sua significância do que os pares entre níveis mais próximos, como o par A4-A2 ou A3-A2 (Tabela 2). Os trabalhos de Nascimento e Sano (2010) e Hermuche e Sano (2011) encontraram padrão semelhante de resultados. A banda 5 foi a que mais se destacou em termos de diferenciação entre as classes após a acumulação das quatro datas ($t = 6,32$, $gl = 10$, $p = 0,0007$) e, analisando todas as bandas, a 5 (infravermelho próximo) e 6 (SWIR 1) foram as que tiveram um maior número de diferenças estatisticamente significativas entre os seis pares de níveis de acumulação. Nascimento e Sano (2010) também encontraram resultado semelhante para a banda do infravermelho médio, e isto se deve provavelmente ao fato de serem estas as bandas que melhor caracterizam a resposta espectral de alvos vegetados, como os deste estudo. As classes “campo” e “várzea” demonstraram, tanto em termos de reflectância individual, quanto acumulada, ter respostas espectrais parecidas, o que denota a dificuldade de diferenciação desses alvos e serve de alerta para posteriores estudos que almejem diferenciá-los com mais precisão. Assim como também sugerido por Nunes e Roig (2015) para a relação entre as classes “campo” e “savana”, as várzeas, apesar de suas particularidades hídricas, podem se comportar espectralmente tal como uma vegetação campestre, especialmente em estações secas, que são bem definidas em função da sazonalidade climática do Cerrado.

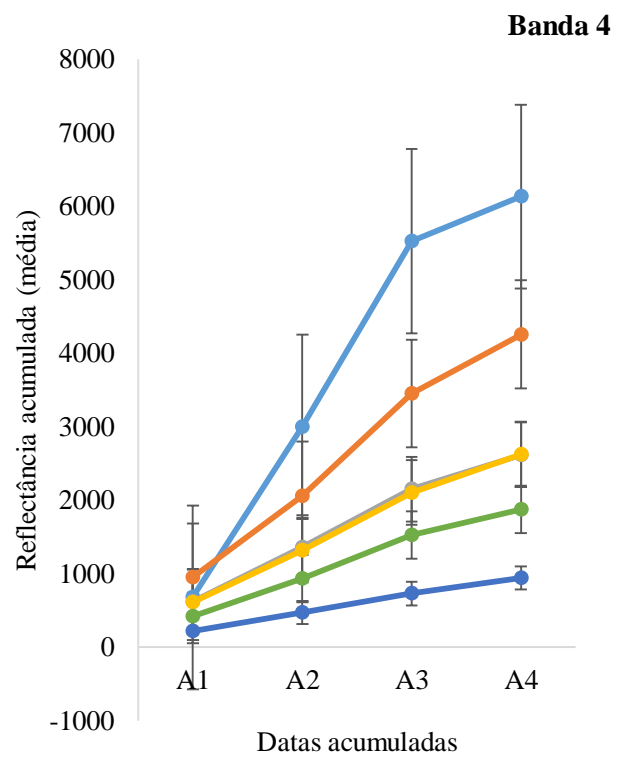
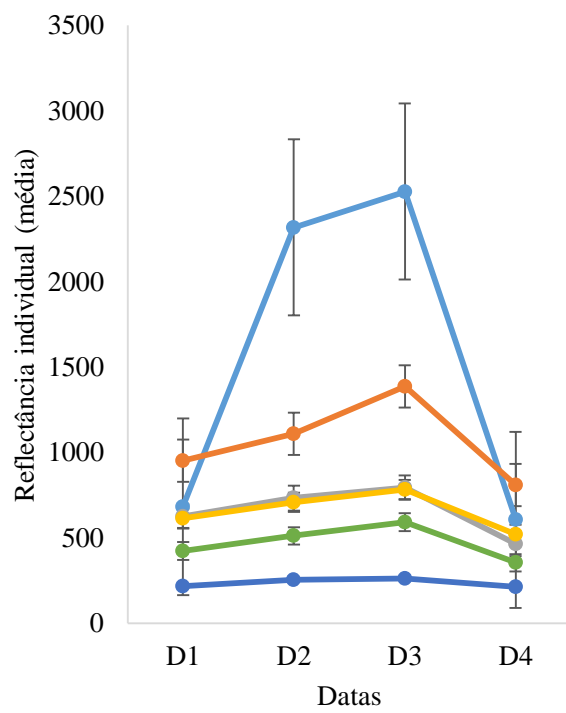
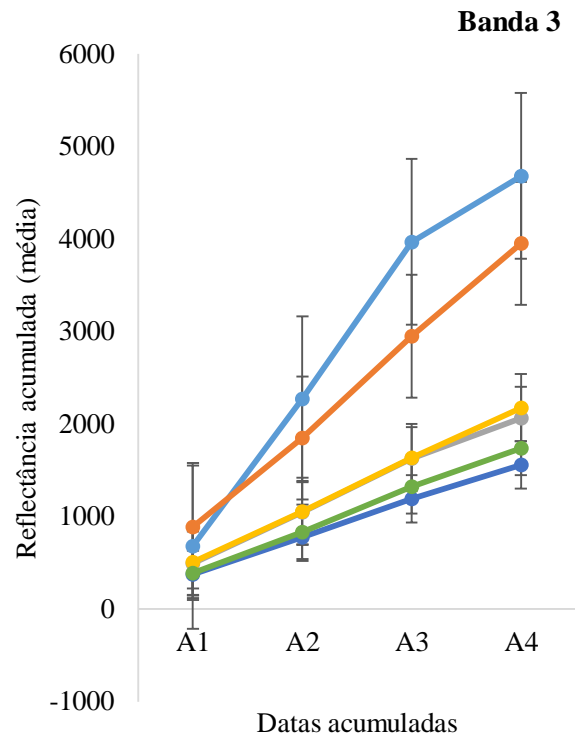
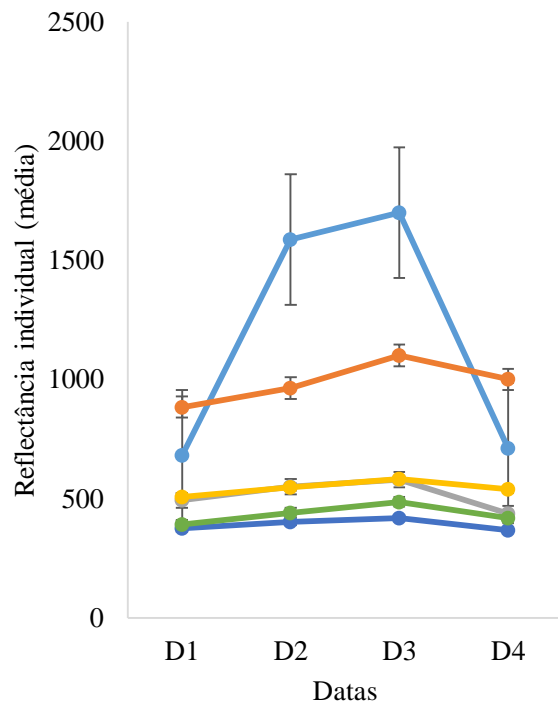


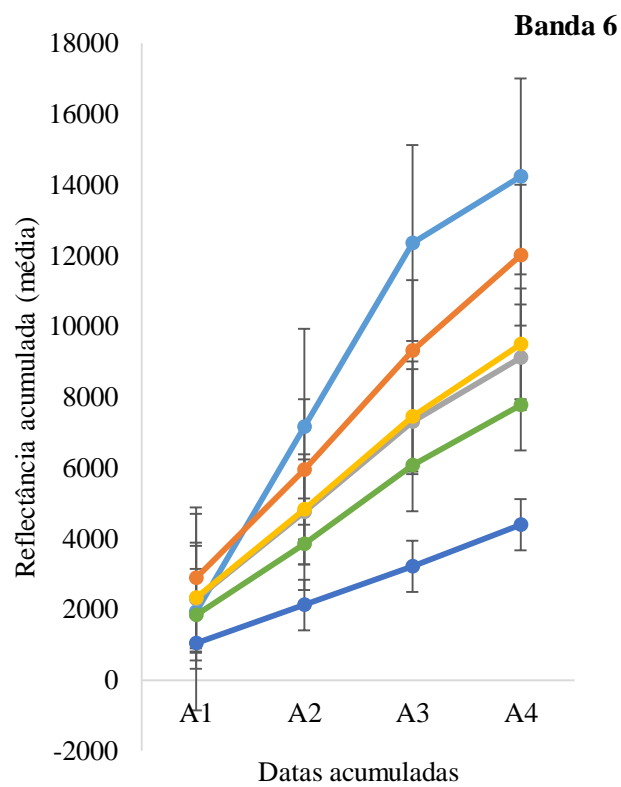
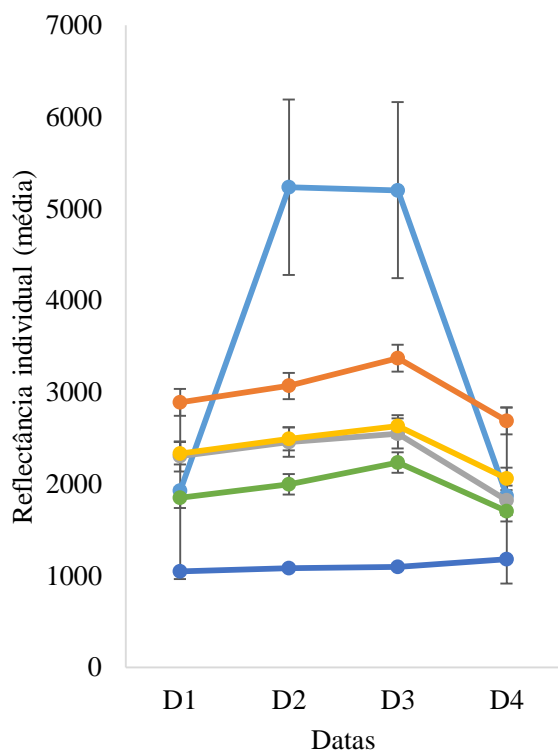
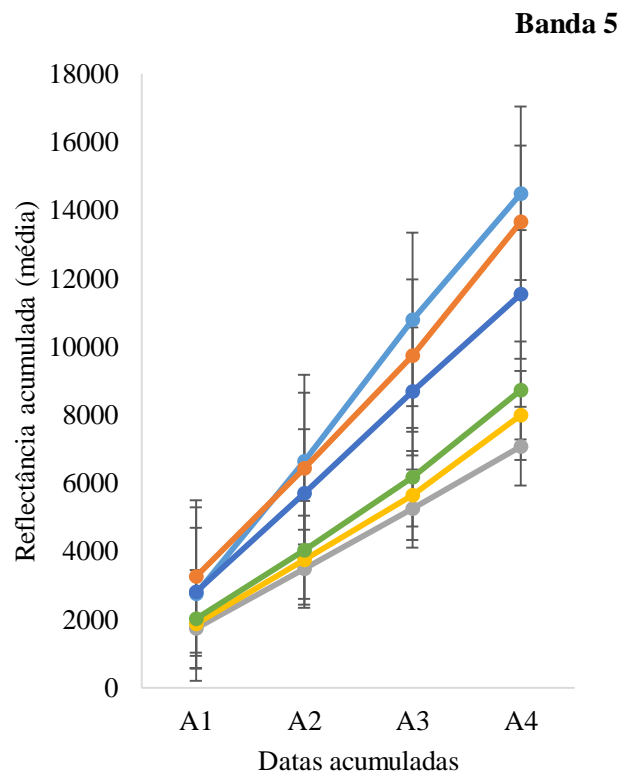
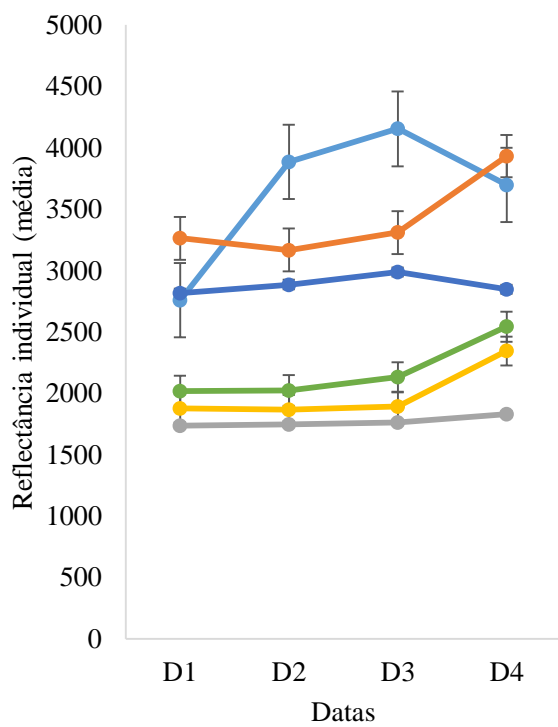
Banda 1



Banda 2







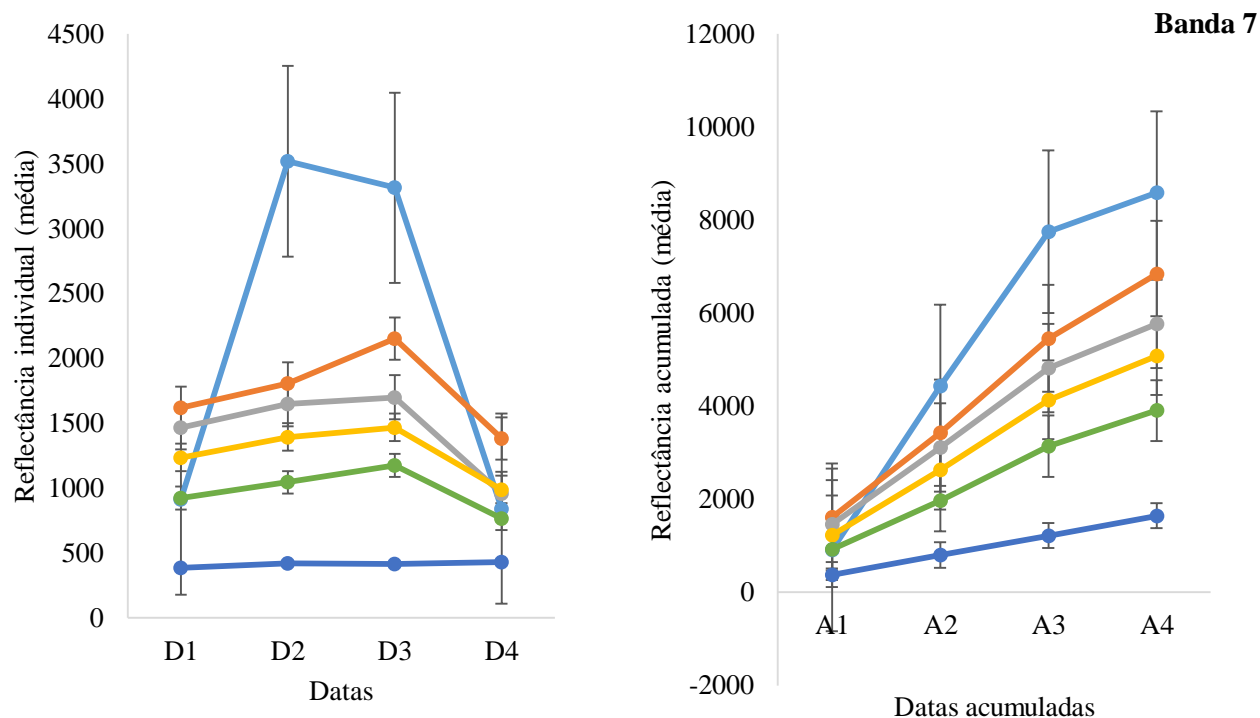


Figura 2. Reflectâncias individuais e acumuladas (com erro padrão) das bandas 1 a 7 do Landsat 8 para seis classes de uso e cobertura da terra do Parque Nacional das Emas e adjacências - Goiás, oriundas de quatro cenas (D1 = 18/07/13, D2 = 03/08/13, D3 = 19/08/13, D4 = 11/02/14), cujos somatórios são representados por A1 = D1, A2 = D1+D2, A3 = D1+D2+D3, A4 = D1+D2+D3+D4.

Tabela 2. Comparações, por banda, das reflectância acumuladas entre os níveis de acumulação das datas (A1, A2, A3 e A4) e entre as seis classes (campo, savana, floresta, várzea, agricultura e pecuária) (ANOVA *two-way* sem repetição), e entre os pares dos níveis de acumulação das datas (Teste t uni-caudal pareado), das cenas da imagem Landsat 8 do Parque Nacional das Emas e adjacências - Goiás. Legenda: A1 = D1, A2 = D1+D2, A3 = D1+D2+D3, A4 = D1+D2+D3+D4.

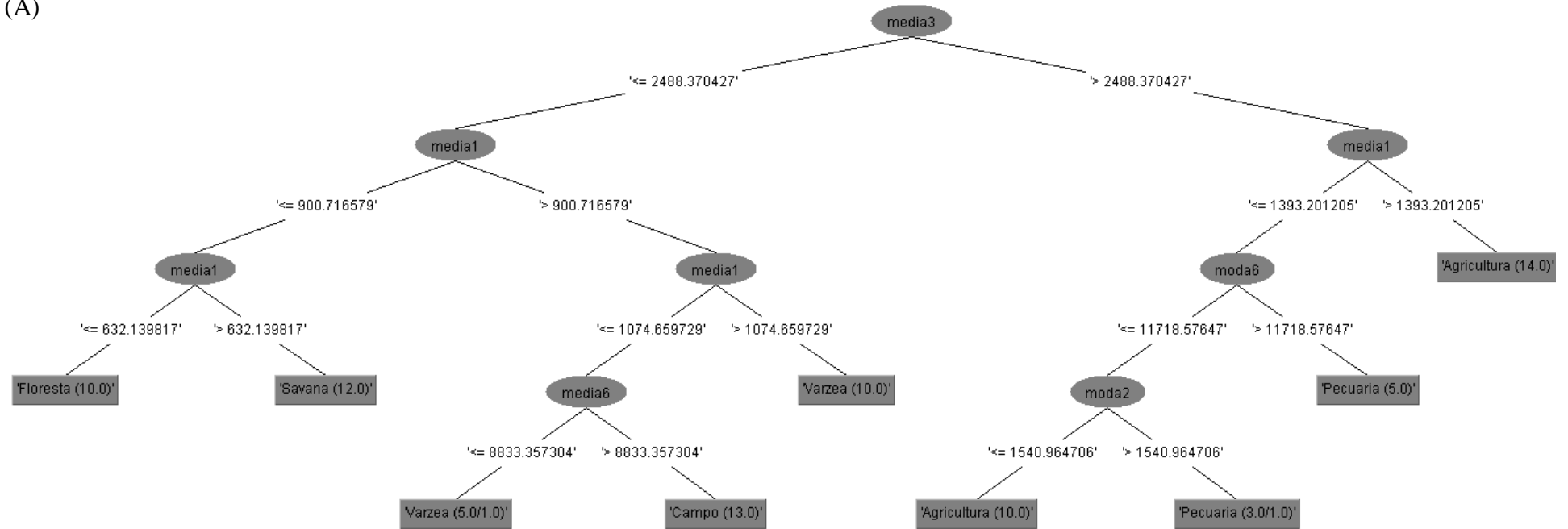
		Bandas						
		1	2	3	4	5	6	7
		ANOVA <i>two-way</i>						
Classes	F	9,30	9,19	9,46	8,71	11,10	9,32	9,20
	p	0,0003	0,0003	0,0003	0,0004	0,0001	0,0003	0,0003
Datas	F	18,01	16,27	19,78	12,14	59,82	32,31	20,67
	p	0,0000	0,0000	0,0000	0,0002	0,0000	0,0000	0,0000
		Teste t						
A4 – A3	t	0,68	0,64	0,83	0,48	1,80	1,00	0,66
	p	0,2530	0,2657	0,2119	0,3198	0,0521	0,1688	0,2598
A4 – A2	t	2,23	2,13	2,37	1,84	4,00	3,03	2,33
	p	0,0302	0,0351	0,0245	0,0534	0,0025	0,0094	0,0238
A4 – A1	t	3,89	3,71	3,99	3,26	6,32	5,27	4,23
	p	0,0057	0,0068	0,0051	0,0111	0,0007	0,0016	0,0040
A3 – A2	t	1,53	1,47	1,57	1,35	2,42	1,97	1,63
	p	0,0811	0,0891	0,0769	0,1067	0,0206	0,0418	0,0702
A3 – A1	t	3,33	3,18	3,40	2,85	5,37	4,34	3,63
	p	0,0103	0,0122	0,0095	0,0178	0,0008	0,0036	0,0075
A2 – A1	t	2,81	2,71	2,86	2,47	4,12	3,62	3,02
	p	0,0151	0,0175	0,0142	0,0239	0,0022	0,0055	0,0116

Graus de liberdade: ANOVA *two-way* - Classes = 5; ANOVA *two-way* - Datas = 3; testes t uni-caudais pareados = 10.
Negrito: significância estatística.

Mineração de dados e classificação orientada a objeto

As figuras 3 e 4, bem como a tabela 3, exemplificam, respectivamente, as árvores de decisão das associações de atributos, suas matrizes de confusão e as estatísticas da mineração de dados. Os resultados da mineração de dados da associação de “todos os atributos juntos” foram considerados, *a priori*, mais adequados para subsidiar a classificação da paisagem de estudo, uma vez que apresentaram, além dos melhores valores de amostras classificadas corretamente e incorretamente e do índice *Kappa*, o menor tamanho de árvore e de número de ramos (Figura 3A; Tabela 3). Entretanto, durante a classificação, o programa InterIMAGE não foi incapaz de processar as regras de decisão da classe “pecuária” (Figura 6B), tornando inabilitado seu nó na rede semântica (Figura 5B) e ausente esta classe no respectivo mapa temático (Figura 7B), deixando certos pixels não classificados (canto esquerdo inferior da figura 7B). Sendo assim, executamos subsequentemente a classificação baseada na associação de atributos “média + moda”, que não apresentou nenhum tipo de impedimento no processamento das regras de decisão (Figura 6A). O minerador de dados utilizou com frequência, em ambas as árvores de decisão, bandas com grande variabilidade, como a banda 1, e menos frequentemente as bandas 4 e 5, que caracterizam bem a resposta espectral da vegetação.

(A)



(B)

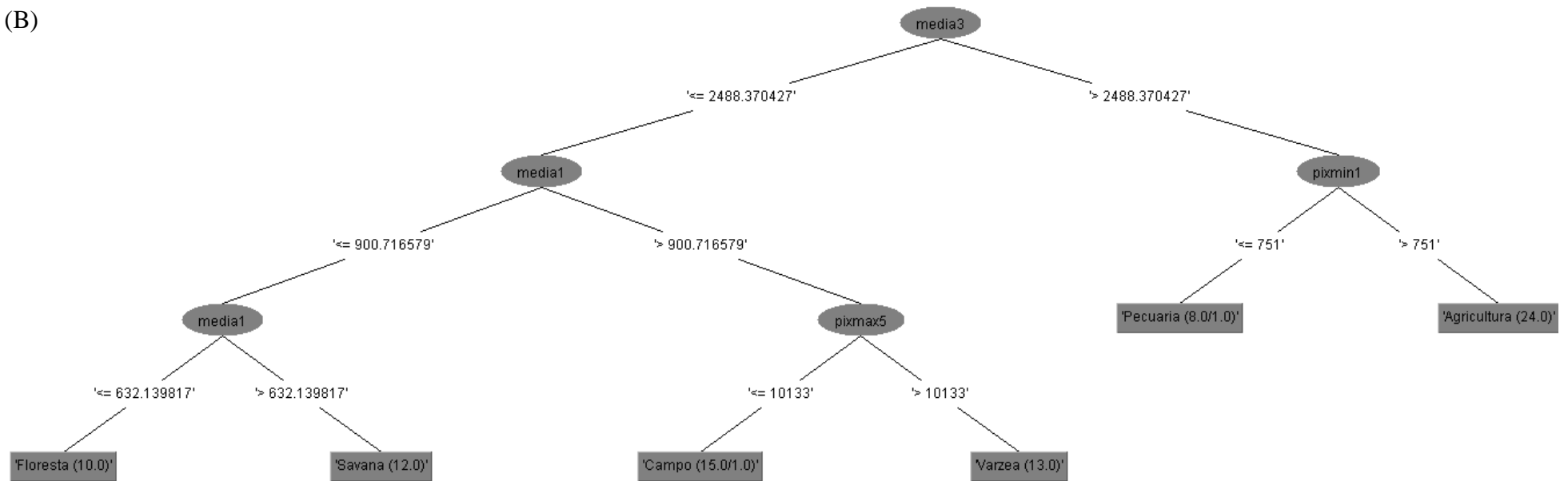


Figura 3. Árvores de decisão geradas pela mineração de dados, operada sobre a associação dos atributos espectrais “média + moda” (A) e “todos os atributos juntos” (média, pixel máximo, pixel mínimo, amplitude, moda, soma) (B), provenientes das amostras de segmentos coletados por classe de uso e cobertura da terra na imagem com reflectância acumulada do Parque Nacional das Emas e adjacências – Goiás.

(A)	(B)
a b c d e f <-- classified as	a b c d e f <-- classified as
24 0 0 1 0 0 a = Agricultura	24 0 0 1 0 0 a = Agricultura
0 13 0 0 0 1 b = Campo	0 14 0 0 0 0 b = Campo
0 0 10 0 0 0 c = Floresta	0 0 10 0 0 0 c = Floresta
0 0 0 7 0 0 d = Pecuária	0 0 0 7 0 0 d = Pecuária
0 0 0 0 12 0 e = Savana	0 0 0 0 12 0 e = Savana
0 0 0 0 0 14 f = Varzea	0 1 0 0 0 13 f = Varzea

Figura 4. Matrizes de confusão das árvores de decisão geradas pela mineração de dados, operada sobre a associação dos atributos espectrais “média + moda” (A) e “todos os atributos juntos” (média, pixel máximo, pixel mínimo, amplitude, moda, soma) (B), provenientes das amostras de segmentos coletados por classe de uso e cobertura da terra na imagem com reflectância acumulada do Parque Nacional das Emas e adjacências – Goiás.

Tabela 3. Estatística oriundas das árvores de decisão geradas pela mineração de dados, operada sobre cada atributo e sobre associações de atributos, provenientes das amostras de segmentos coletados por classe de uso e cobertura da terra na imagem com reflectância acumulada do Parque Nacional das Emas e adjacências – Goiás.

Classe	Número de caminhos que culminam nas classes						Média + Moda	Todos atributos
	Média	Pixmax	Pixmin	Amplitude	Moda	Soma		
Agricultura (25)	3	3	2	7	3	3	2	1
Pecuária (7)	1	1	1	2	2	1	2	1
Floresta (10)	1	3	1	2	1	2	1	1
Campo (14)	1	2	1	1	1	1	1	1
Varzea (14)	2	3	3	4	2	3	2	1
Savana (12)	1	3	1	1	1	3	1	1
Número de ramos	9	15	9	17	10	13	9	6
Tamanho da árvore	17	29	17	33	19	25	17	11
Amostras classificadas corretamente	80 (97,5%)	77 (94,0%)	77 (94,0%)	70 (85,4%)	79 (96,4%)	69 (84,2%)	80 (97,5%)	80 (97,5%)
Amostras classificadas incorretamente	2 (2,5%)	5 (6,0%)	5 (6,0%)	12 (14,6%)	3 (3,6%)	13 (15,8%)	2 (2,5%)	2 (2,5%)
Kappa	0,969	0,923	0,923	0,817	0,954	0,803	0,969	0,969

Classe () = número de amostras coletadas, total de 82. Média: valor médio de pixels; Pixmax: valor de pixel máximo por banda; Pixmin: valor de pixel mínimo por banda; Amplitude: diferença entre os valores máximos e mínimos de pixel por banda; Moda: valor mais frequente de pixels por banda; Soma: somatório dos valores de pixel por banda; Média + moda: atributos média e moda associados; Todos atributos: todos atributos associados. Negrito: valores mais expressivos.

(A)



(B)

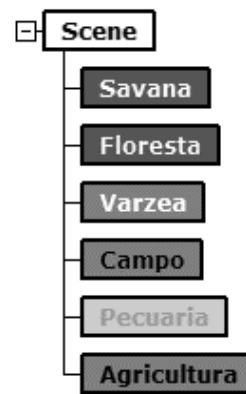
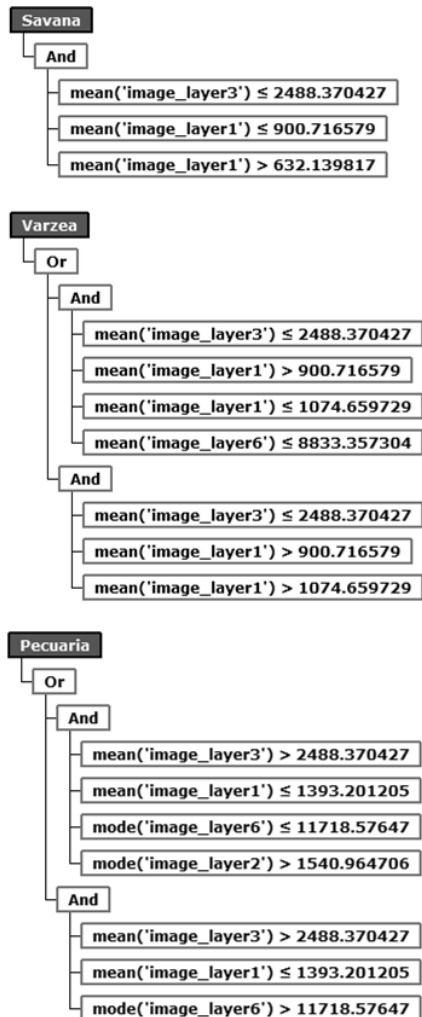
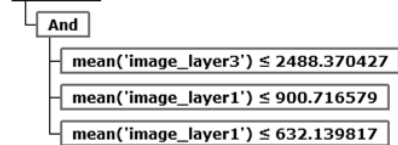


Figura 5. Redes semânticas das classificações baseadas nas árvores de decisão das associações dos atributos espectrais “média + moda” (A) e “todos os atributos juntos” (média, pixel máximo, pixel mínimo, amplitude, moda, soma) (B), provenientes das amostras de segmentos coletados por classe de uso e cobertura da terra na imagem com reflectância acumulada do Parque Nacional das Emas e adjacências – Goiás. Em (B), o programa não foi capaz de classificar a classe “pecuária”.

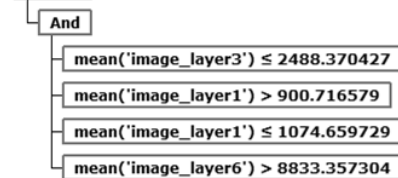
(A)



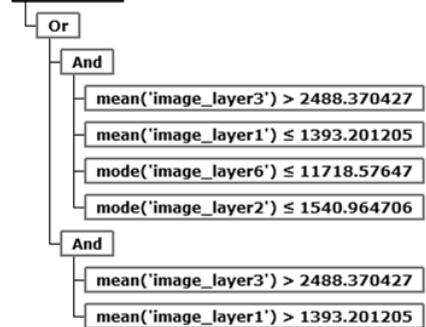
Floresta



Campo



Agricultura



(B)

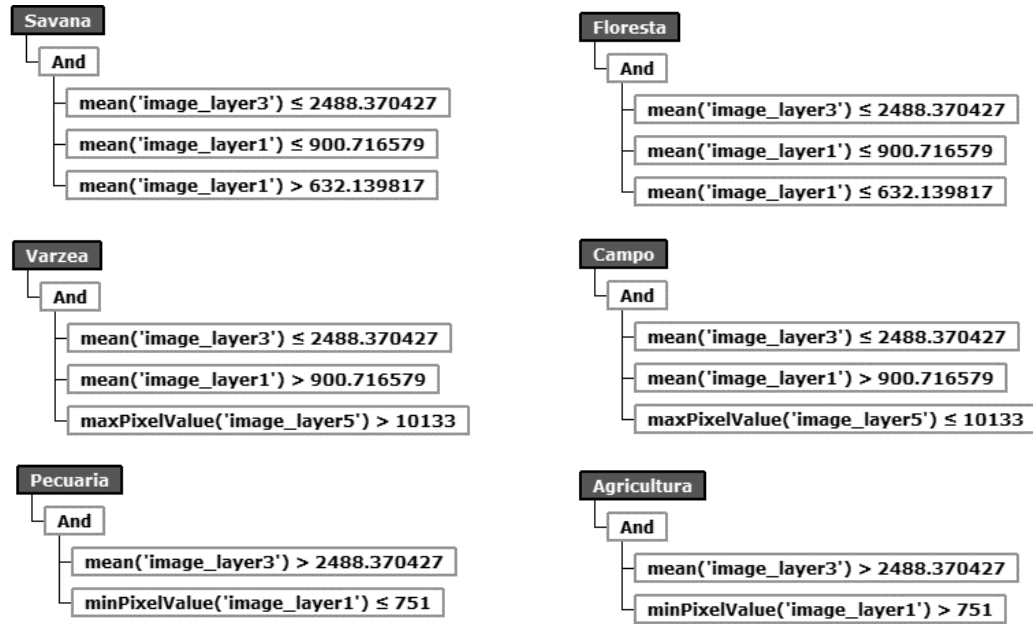


Figura 6. Regras de decisão *TopDown* construídas a partir das árvores de decisão baseadas nas associações dos atributos “média + moda” (A) e “todos os atributos juntos” (média, pixel máximo, pixel mínimo, amplitude, moda, soma) (B), provenientes das amostras de segmentos coletados por classe de uso e cobertura da terra na imagem com reflectância acumulada do Parque Nacional das Emas e adjacências – Goiás.

Verificação de qualidade da classificação

A classificação baseada na associação de atributos “média + moda” apresentou boa acurácia (índice de Exatidão Global de 69%, índice *Kappa* de 58% e índice TAU de 63%) (Tabela 4 e 5). O mapa temático (Figura 7A) indicou que as classes “savana”, “floresta” e “campo” apresentaram razoável concordância na classificação. No entanto, a classe “várzea” se confundiu, em muitas regiões, com a classe “campo”, talvez em função da sua semelhança espectral observada nos gráficos de reflectância, e com áreas de estradas de terra. Este último caso denota, particularmente, uma falha na legenda definida, uma vez que deveríamos ter considerado uma classe denominada “outros”. A classe “agricultura” também se mostrou relativamente bem classificada, preenchendo a maior parte de área por ela de fato ocupada na paisagem; no entanto, algumas regiões agrícolas foram atribuídas como áreas de “pecuária”. Já as poucas áreas realmente dominadas por pecuária foram corretamente classificadas.

Tabela 4. Matriz de confusão contendo os dados de referência e os dados classificados, para fins de verificação da qualidade da classificação *TopDown* baseada na associação de atributos “média + moda”, na imagem com reflectância acumulada do Parque Nacional das Emas e adjacências – Goiás.

		Dados de referência						
Dados classificados		Agricultura	Campo	Floresta	Pecuária	Savana	Várzea	Total
	Agricultura	48	1	1	2	1	1	54
	Campo	2	10	0	0	3	0	15
	Floresta	0	0	7	0	0	0	7
	Pecuária	5	0	0	2	0	1	8
	Savana	2	1	1	0	13	4	21
	Várzea	0	11	0	0	2	8	21
	Total	57	23	9	4	19	14	126

Tabela 5. Área ocupada por cada classe e índices de verificação da qualidade da classificação *TopDown* baseada na associação de atributos “média + moda”, na imagem com reflectância acumulada do Parque Nacional das Emas e adjacências – Goiás.

Classe	Área	Porcentagem
Agricultura	511908300	28,89
Campo	582691500	32,89
Floresta	46654200	2,63
Pecuária	102640500	5,79
Savana	205002900	11,57
Várzea	322561800	18,20
Total	1771459200	100
Índices da verificação de qualidade		
Exatidão Global (Po)	0,698	
Kappa	0,589	
TAU	0,638	

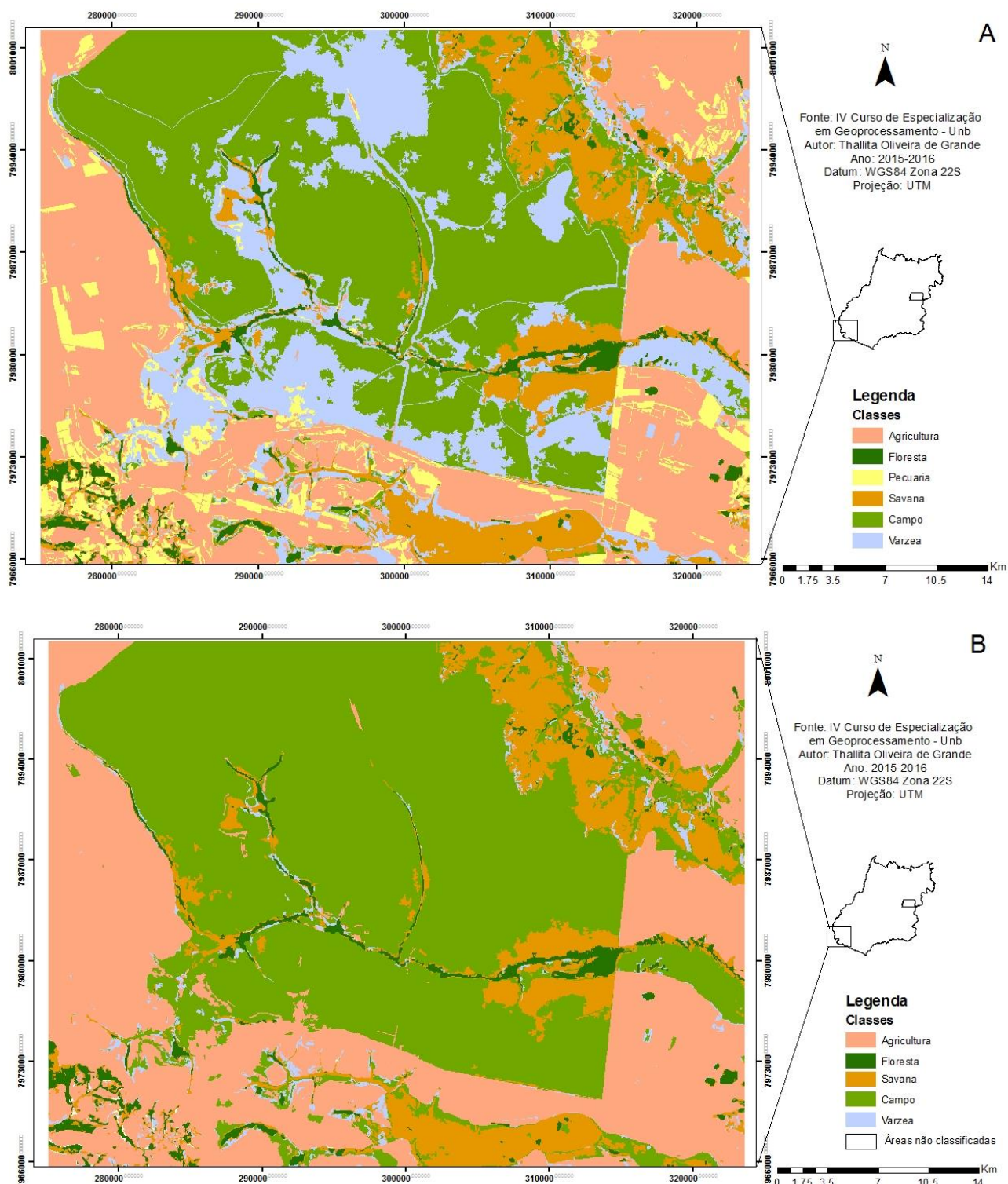


Figura 7. Mapas temáticos das classificações orientadas a objeto baseadas nas associações dos atributos “média + moda” (A) e “todos os atributos juntos” (B), da imagem do Parque Nacional das Emas e adjacências – Goiás.

Testes *a posteriori*

Após os primeiros resultados obtidos, repetimos, *a posteriori*, a metodologia anteriormente descrita na tentativa de melhorar a eficácia da discriminação das classes. Como base nos parâmetros usados na tabela 1 para a classificação, segmentamos a imagem, escolhemos 853 amostras, e

extraímos três grupos de atributos (Tabela 6). Justificamos a inclusão dos atributos de forma e vizinhança em função do fato de que as classes nativas e antrópicas da paisagem do Cerrado são geométrica e estruturalmente bem caracterizadas (ex.: pivôs, campos agrícolas, áreas florestais), elementos que podem ser usados a favor da distinção entre classes. A partir dos resultados da mineração de dados (Tabela 6), escolhemos o terceiro grupo e realizamos a classificação. Entretanto, as árvores de decisão oriundas desta associação eram demasiado complexas e o InterIMAGE não foi capaz de classificar todas as classes, apenas “savana” e “floresta”.

Tabela 6. Estatística oriundas das árvores de decisão geradas pela mineração de dados, operada sobre as associações de atributos, provenientes das amostras de segmentos coletados, nos testes *a posteriori*, por classe de uso e cobertura da terra na imagem com reflectância acumulada do Parque Nacional das Emas e adjacências – Goiás.

Classe	Número de caminhos que culminam nas classes		
	Atributos espectrais	Média + moda	Atributos espectrais + forma + vizinhança
Agricultura (189)	11	12	7
Pecuária (105)	5	8	2
Floresta (125)	1	2	1
Campo (100)	3	4	3
Várzea (118)	7	6	6
Savana (116)	3	7	2
Outros (100)	12	13	11
Número de ramos	42	52	32
Tamanho da árvore	83	103	63
Amostras classificadas corretamente	834 (97,7%)	832 (97,5%)	834 (97,7%)
Amostras classificadas incorretamente	19 (2,3%)	21 (2,5%)	19 (2,3%)
Kappa	0,973	0,971	0,973

Classe () = número de amostras coletadas, total de 853. Atributos espectrais (Média: valor médio de pixels; Pixmax: valor de pixel máximo por banda; Pixmin: valor de pixel mínimo por banda; Amplitude: diferença entre os valores máximos e mínimos de pixel por banda; Moda: valor mais frequente de pixels por banda; Soma: somatório dos valores de pixel por banda; Amplitude: diferença entre os valores máximos e mínimos de pixel de um objeto; Variância: grau de dispersão dados numéricos em torno da média, mas na escala de valores de dados original; Brilho: brilho de um objeto na imagem). Média + moda: atributos média e moda associados. Atributos de forma (Compacidade: compacidade do polígono; Densidade: razão entra a área e o rádio do polígono; Índice de forma: índice de forma do polígono; Razão perímetro/área: razão entre o perímetro e a área do polígono; Razão comprimento/largura: razão entre o comprimento e a largura do polígono). Atributos de vizinhança (Existência_vizinho: existência de um polígono vizinho pertencente a uma classe selecionada, dentro de um intervalo (em pixels) em torno do polígono alvo; Número_vizinho: número de polígonos vizinhos pertencente a uma classe selecionada, dentro de um intervalo (em pixels) em torno do polígono alvo; Área_vizinho: área preenchida por polígonos pertencentes a uma classe selecionada, dentro de um intervalo (em pixels) em torno do polígono alvo, dividido pela área total de polígonos no interior do intervalo. Atributos espectrais + forma + vizinhança: atributos espectrais, de forma e de vizinhança associados. Negrito: valores mais expressivos.

Demais considerações

Diante de todos os resultados, podemos observar que não necessariamente a árvore de decisão com menor número de ramos e tamanho gera a regra de decisão que melhor subsidia a classificação. Neste estudo, percebemos o contrário, em que a regra de decisão oriunda da mineração de dados sobre a associação de atributos “média + moda” ofereceu a melhor classificação. Já sobre a associação de “todos atributos juntos”, matematicamente mais simples, a classificação não conseguiu ser finalizada. O diferencial pode ter sido o fato da primeira associação englobar um menor número de atributos. Como lembra Oliveira et al. (2007), podemos acreditar que quanto mais atributos usarmos para representar uma classe, maior seria a eficácia da classificação; porém, isto não se faz verdadeiro uma vez que o aumento do número de atributos acarreta no aumento da dimensionalidade dos dados e na diminuição da acurácia dos resultados, quando o número de amostras se mantém constante. Os testes *a posteriori* confirmaram este apontamento, uma vez que apesar de termos coletado um maior número de amostras numa escala de segmentação mais refinada, e do pequeno aumento absoluto do índice *Kappa* em relação aos resultados das primeiras minerações, a execução da classificação foi inviável, muito provavelmente em função do aumento do número de atributos.

Outra questão que devemos pontuar é o fato de que as regras de decisão obtidas com base na análise de uma paisagem podem não se aplicar com eficácia a outra paisagem do mesmo bioma, inclusive dominadas pelas mesmas classes de uso e cobertura da terra. Isto denota certa falta de generalidade desta técnica; logo, propomos que estudos futuros testem as regras de decisão obtidas em uma paisagem em outras similares, ou refinem as regras de decisão de cada classe, no intuito de encontrar um denominador comum entre as mesmas classes, em paisagens diferentes. A importância de tornar as regras de decisão replicáveis a fim de aumentar a aplicabilidade da técnica também é algo sugerido por Seabra et al. (2013), no mapeamento de uma paisagem carioca utilizando, igualmente, a mineração de dados e classificação baseada em objetos.

O aumento sutil da complexidade da árvore e regra de decisão da associação de “todos os atributos” para a associação “média + moda” não influenciou o desempenho da classificação. Porém, o considerável aumento da complexidade da árvore e regras de decisão, demonstrado por meio dos testes *a posteriori*, culminou na incapacidade de processamento pelo InterIMAGE. Isto nos leva a questionar qual o limiar dessa capacidade no contexto computacional do programa. Ruiz et al. (2014), discutindo sobre árvores de decisão e a densidade mínima de amostras no mapeamento da cobertura da terra, afirmam que o aumento do número de amostras acarreta no aumento da complexidade das árvores de decisão, dificultando as classificações. Outro ponto a ser ponderado, no que diz respeito às falhas no processamento da classificação que observamos, é se a área da imagem utilizada era demasiadamente extensa. Apesar de estar dentro dos limites suportados pelo programa, de acordo com seus criadores (Laboratório de Visão Computacional 2012), talvez o aumento da complexidade

das regras de decisão inseridas possa relativizar o tamanho da imagem utilizada. Nascimento et al. (2013) também enfrentaram problemas no que diz respeito a limitação no tamanho das imagens que podem ser inseridas no sistema e interrupções abruptas no seu funcionamento, e indicam a necessidade de amadurecimento computacional do programa. Assim como o presente estudo, Junior et al. (2014), comparando a classificação supervisionada e a orientada a objeto em imagens Landsat 8, também pontuam barreiras de processamento no uso do programa, como a incapacidade de lidar com um número grande de segmentos e a falta de automação.

Como alternativas para a melhoria dos resultados que obtivemos, uma possibilidade seria, segundo o trabalho de Carvalho et al. (2014), construir a rede semântica subdividindo os nós em nós filhos, ou refinar os resultados buscando manualmente os limiares entre as classes, por meio da análise de gráficos de histogramas fornecidos pela ferramenta *Analysis Manager*, utilizada geralmente com a ferramenta *Debug Mode* ativa e com *Breakpoint* selecionado para a classificação *TopDown*. Outra possibilidade seria testar diferentes tipos de operadores para cada nó da rede semântica, embora esta alternativa possa acarretar em custos operacionais (Novack 2009; Carvalho et al. 2014).

CONCLUSÃO

Finalmente, podemos dizer que a associação destas três técnicas (reflectância acumulada, mineração de dados e classificação orientada a objeto) têm potencial na diferenciação dos alvos vegetados nativos e antrópicos na região do Parque Nacional das Emas e adjacências, podendo, mediante estudos futuros, ser aplicada para o bioma todo. Especialmente a reflectância acumulada e a mineração de dados demonstraram ser técnicas perfeitamente factíveis e profícuas. Contudo, questionamos a exequibilidade da classificação orientada a objeto no programa InterIMAGE e sua capacidade computacional. Mais especificamente, nosso questionamento tange os limiares que regem a escolha dos atributos elencados para caracterizar a formações nativas e antrópicas das vegetações do Cerrado, a complexidade das árvores de decisão e, conseqüentemente, das regras de decisão implantadas no classificador para fins da separação temática dos alvos. O aprofundamento destes aspectos pode ser alvo de investigações científicas futuras.

REFERÊNCIAS BIBLIOGRÁFICAS

- Antunes, R.R., Bias, E. de S., Brites, R. S., da Costa, G. A. O. P. 2015. Desenvolvimento de técnica para monitoramento do cadastro urbano baseado na classificação orientada a objetos. Estudo de caso: município de Goianésia-go. *Revista Brasileira de Cartografia*, 67/2: 357-372.
- Baatz, M. e Schäpe, A. 2000. Multiresolution segmentation: an optimization approach for high quality multi-scale image segmentation. In: XII Angewandte Geographische Informationsverarbeitung, Agit Symposium, Salzburg, Áustria, p. 12-23.

- Blaschke, T. 2010. Object based image analysis for remote sensing. *ISPRS journal of photogrammetry and remote sensing*, 65: 2-16.
- Boardman, J. W. e Kruse, F. A. 1994. Automated spectral analysis: A geologic example using AVIRIS data, north Grapevine Mountains, Nevada. In: *Proceedings of 10th Thematic Conference on Geologic Remote Sensing*, Environmental Research Institute of Michigan, Ann Arbor, MI, p. 407-418.
- Brannstrom, C., Jepson, W., Filippi, A. M., Redo, D., Xu, Z., Ganesh, S. 2008. Land change in the Brazilian Savanna (Cerrado), 1986–2002: comparative analysis and implications for land-use policy. *Land Use Policy*, 25: 579-595.
- Carvalho Júnior, O. A., Carvalho, A. P. F, Guimarães, R. F., Meneses, P. R., Shimabukuro, Y. E. 2003. Mistura espectral: (I) Detecção dos membros finais utilizando a geometria do simplex. *Espaço & Geografia*, 6: 147-173.
- Carvalho Júnior, O. A., Ribeiro, M. N. C., Martins, E. S. 2007. Análise de dados hiperespectrais em mina de manganês laterítico, São João da Aliança (GO). *Revista Brasileira de Geofísica*, 25: 199-212.
- Carvalho, A. A. A. A. et al. 2014. Análise de classificação de imagem Rapideye por meio dos softwares Interimage e Spring. *Revista Brasileira de Cartografia*, ?-?
- Carvalho, A. P. F., Carvalho Júnior, O. A., Guimarães, R. F., Martins, E. S., Bustamante, M. M. C. 2002. Utilização do classificador *spectral correlation mapper* em imagens TM-Landsat. *Espaço & Geografia*, 5: 219-232.
- Cohen, J. 1960. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20: 37-46.
- Congalton, R. G. e Green, K. 1999. *Assessing the accuracy of remotely sensed data: principles and practices*. New York: Lewis Publishers 137 p.
- Costa, W. S. 2014. Mapeamento de formações campestres nativas e de pastagens cultivadas no cerrado brasileiro utilizando mineração de dados. Dissertação de Mestrado do Curso de Pós-Graduação em Computação Aplicada, INPE, São José dos Campos. 85p.
- DeFries, R. S. e Chan, J. 2000. Multiple criteria for evaluating machine-learning algorithms for land-cover classification from satellite data. *Remote Sensing of Environment*, 74: 503–515.
- Faleiro, F. V., Machado, R. B., Loyola, R. D. 2013. Defining spatial conservation priorities in the face of land-use and climate change. *Biological Conservation*, 158: 248-257.
- Green, A. A., Berman, M., Switzer, P., Craig, M. D. 1988. A transformation for ordering multispectral data in terms of images quality with implications for noise removal. *IEEE Transactions on Geoscience & Remote Sensing*, 26: 65-74.
- Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I. H. 2009. The WEKA data mining software: an update. *ACM SIGKDD explorations newsletter*, 11: 10-18.
- Hellden, U. e Stern, M. 1980. Evaluation of landsat imagery and digital data for monitoring desertification indicators in Tunisia. In: *Proc. 14th. Int. Symp. on Rem. Sens. of Environ.*, p. 1601-1611.
- Hermuche, P. M. e Sano, E. E. 2011. Identificação da floresta estacional decidual no Vão do Paranã, estado de Goiás, a partir da análise da reflectância acumulada de imagens do sensor ETM+/Landsat-7. *Revista Brasileira de Cartografia*, 63/03: 415-425.
- Ibama. 2004. *Plano de manejo do Parque Nacional das Emas*. Brasília. 893p.

- Júnior, E. S. A., de Lima Nascimento, D., dos Santos Galvanin, E. A., da Silva Neves, S. M. A., de Jesus, P. H. H., de Gouveia, R. G. L. 2014. Comparação dos métodos de classificação supervisionada e orientada a objeto em imagens Landsat 8 na Bacia hidrográfica Paraguai/Jauquara–Mato Grosso. In: Anais 5º Simpósio de Geotecnologias no Pantanal, Campo Grande, MS, Embrapa Informática Agropecuária/INPE, p. 839 -847.
- Klecka, W. R. 1980. Discriminant Analysis. SAGE Publications, Inc., Beverly Hills, California, 71p.
- Klink, C. A., Machado, R. B. 2005. Conservation of the Brazilian Cerrado. *Conservation Biology*, 19: 707-713.
- Laboratório de Visão Computacional. Pontifícia Universidade Católica do Rio de Janeiro. 2012. InterIMAGE 1.30 - Manual do usuário. Disponível em: <http://www.lvc.ele.puc-rio.br/projects/interimage/pt-br/documentacao/>
- Liu, Y., Li, M., Mao, L., Xu, F., Huang, S. 2006. Review of remotely sensed imagery classification patterns based on object-oriented image analysis. *Chinese Geographical Science*, 16: 282-288.
- Myers, N., Mittermeier, R. A., Mittermeier, C. G., Fonseca, G. A., Kent, J. 2000. Biodiversity hotspots for conservation priorities. *Nature*, 403: 853-858.
- Nascimento, E. R. P. e Sano, E. E. 2010. Identificação de Cerrado Rupestre por meio de imagens multitemporais do Landsat: proposta metodológica. *Sociedade & Natureza*, 22: 93-106.
- Nascimento, A. F. do, et al. 2013. Classificação da Cobertura da terra, utilizando os programas livres: InterIMAGE, WEKA e QuantumGIS. In: Anais XVI Simpósio Brasileiro de Sensoriamento Remoto - SBSR, Foz do Iguaçu, PR, INPE, p. 2258-2265.
- Novack, T. 2009. Classificação da cobertura da terra e do uso do solo urbano utilizando o sistema InterIMAGE e imagens do sensor Quickbird. INPE-14183-TDI/1095). Dissertação de Mestrado em Sensoriamento Remoto, INPE, São José dos Campos. 214p.
- Nunes, J. F., e Roig, H. L. 2015. Análise e mapeamento do uso e ocupação do solo da bacia do alto do descoberto, df/go, por meio de classificação automática baseada em regras e lógica nebulosa. *Revista Árvore*, 9: 25-36.
- Oliveira, J. A., Dutra. L. V., Rennó. C. D. 2007. Aplicação de métodos de extração e seleção de atributos para classificação de regiões. In: Anais XII Simpósio Brasileiro de Sensoriamento Remoto, Goiânia, GO, INPE, p. 4201-4208.
- Passo, D. P., de Souza Bias, E., Brites, R. S., Costa, G. A. 2013. Uso do sistema InterIMAGE para a identificação de alvos urbanos em imagens do satélite Worldview II. *Revista Brasileira de Cartografia*, 65/6: 1211-1221.
- Pinheiro, E. D. S., e Durigan, G. 2009. Dinâmica espaço-temporal (1962-2006) das fitofisionomias em unidade de conservação do Cerrado no sudeste do Brasil. *Revista Brasileira de Botânica*, 32: 441-454.
- Quinlan, J.R. 1993. C4.5: programs for machine learning. San Francisco: Morgan Kaufmann. 301p.
- Sano, E. E., Rosa, R., Brito, J. L. S., Ferreira, L. G., Bezerra, H. S. 2009. Mapeamento da cobertura vegetal natural e antrópica do bioma Cerrado por meio de imagens Landsat ETM+. In: Anais do Simpósio Brasileiro de Sensoriamento Remoto, Natal, RN, INPE, p. 1199-1206.
- Sano, E. E., Ferreira, L. G., Asner, G. P., Steinke, E.T. 2007. Spatial and temporal probabilities of obtaining cloud-free Landsat images over the Brazilian tropical savanna. *International Journal of Remote Sensing*, 28: 2739-2752.
- Sano, E. E., Rosa, R., Brito, J. L., Ferreira, L.G. 2010. Land cover mapping of the tropical savanna region in Brazil. *Environmental monitoring and assessment*, 166: 113-124.

- Seabra, V. da S., Kalaf, R., Cardoso, P. V., Cruz, C. B. M. 2013. Mineração de dados e classificação baseada em objetos no mapeamento dos estados de médio prazo da paisagem na bacia hidrográfica do Rio São João, RJ. In: Anais XVI Simpósio Brasileiro de Sensoriamento Remoto - SBSR, Foz do Iguaçu, PR, INPE, p. 7770-7777.
- Whitten, I .H., Frank, E. 2005. Data mining – practical machine learning tools and techniques. 2.ed. San Francisco: Morgan Kaufmann. 558p.